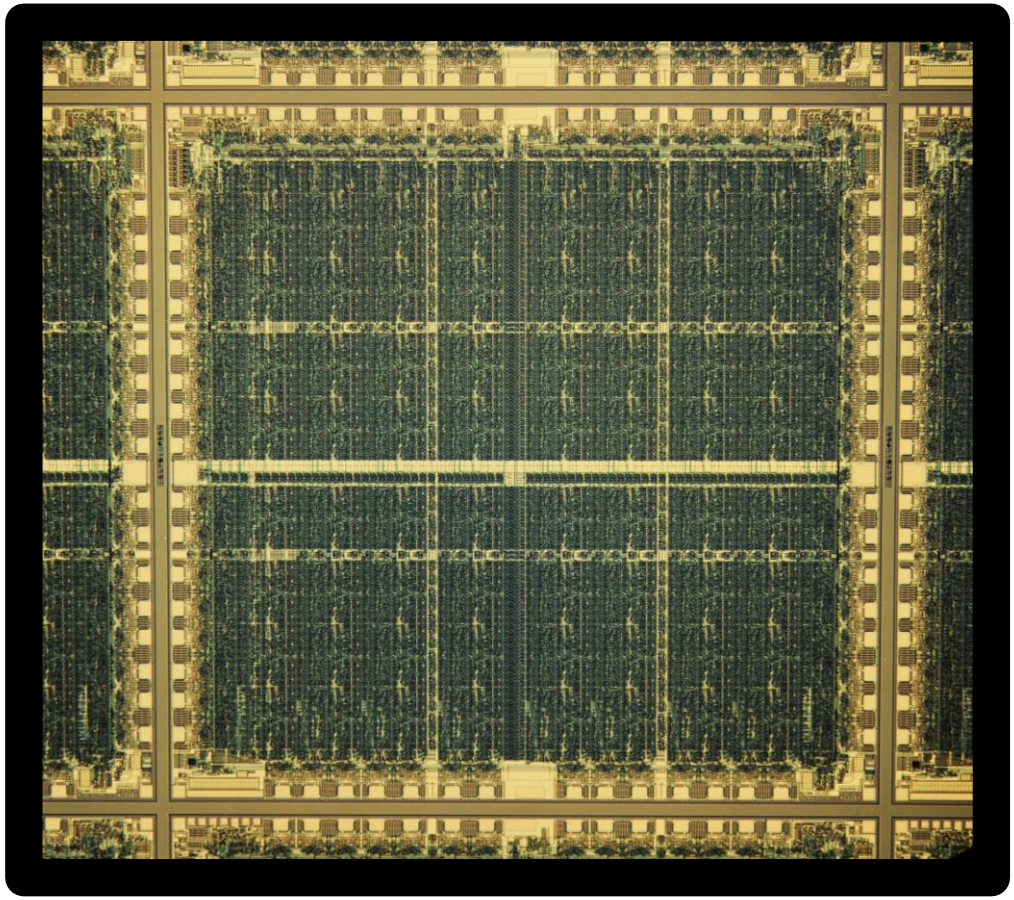
UltraScale+ MPSoC and FPGA Families

Vamsi Boppana, Sagheer Ahmad, Ilya Ganusov, Vinod Kathail
Vidya Rajagopalan, Ralph Wittig

HotChips 27, August 2015

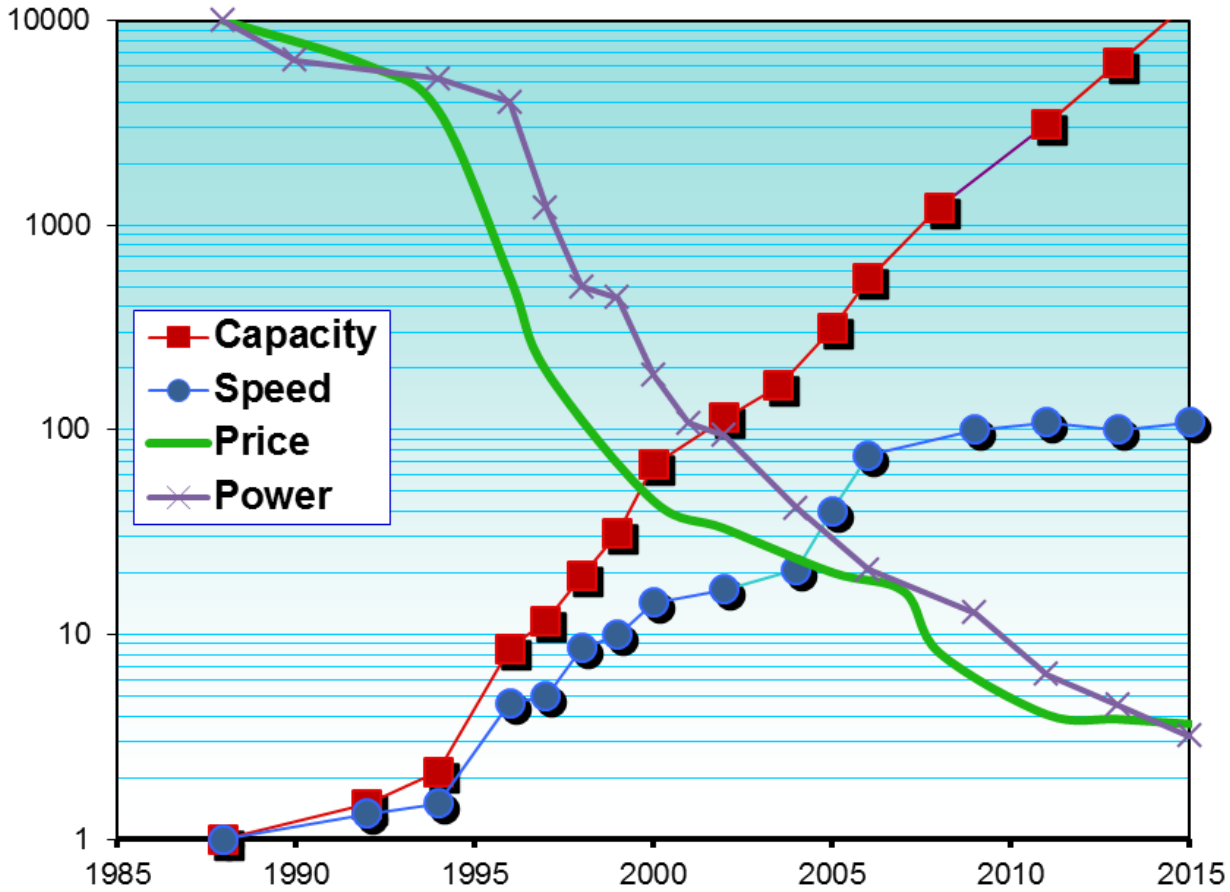
Introducing the XC2064…



The First FPGA (1985)

- 64 flip flops
- 128 3-LUTs
- 58 I/O pins
- 18MHz (toggle)
- 2um 2LM

Since then…

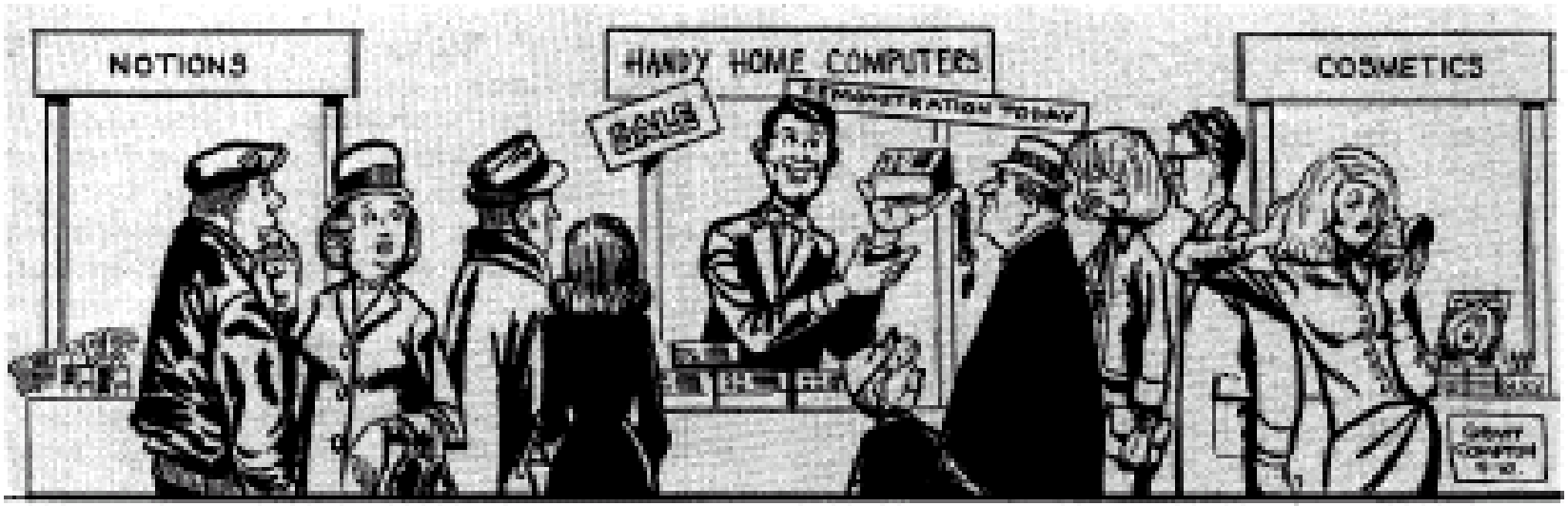


- **10,000x More Logic**
  - Plus Embedded IP
    - Memory
    - Microprocessor
    - DSP
    - Gigabit Serial I/O
- **100x Faster**
- **5000x Lower Power**
- **10,000x Lower Cost**

Three Ages of the FPGAs: Trimberger, S, Proceedings of the IEEE | Vol. 103, No. 3, March 2015

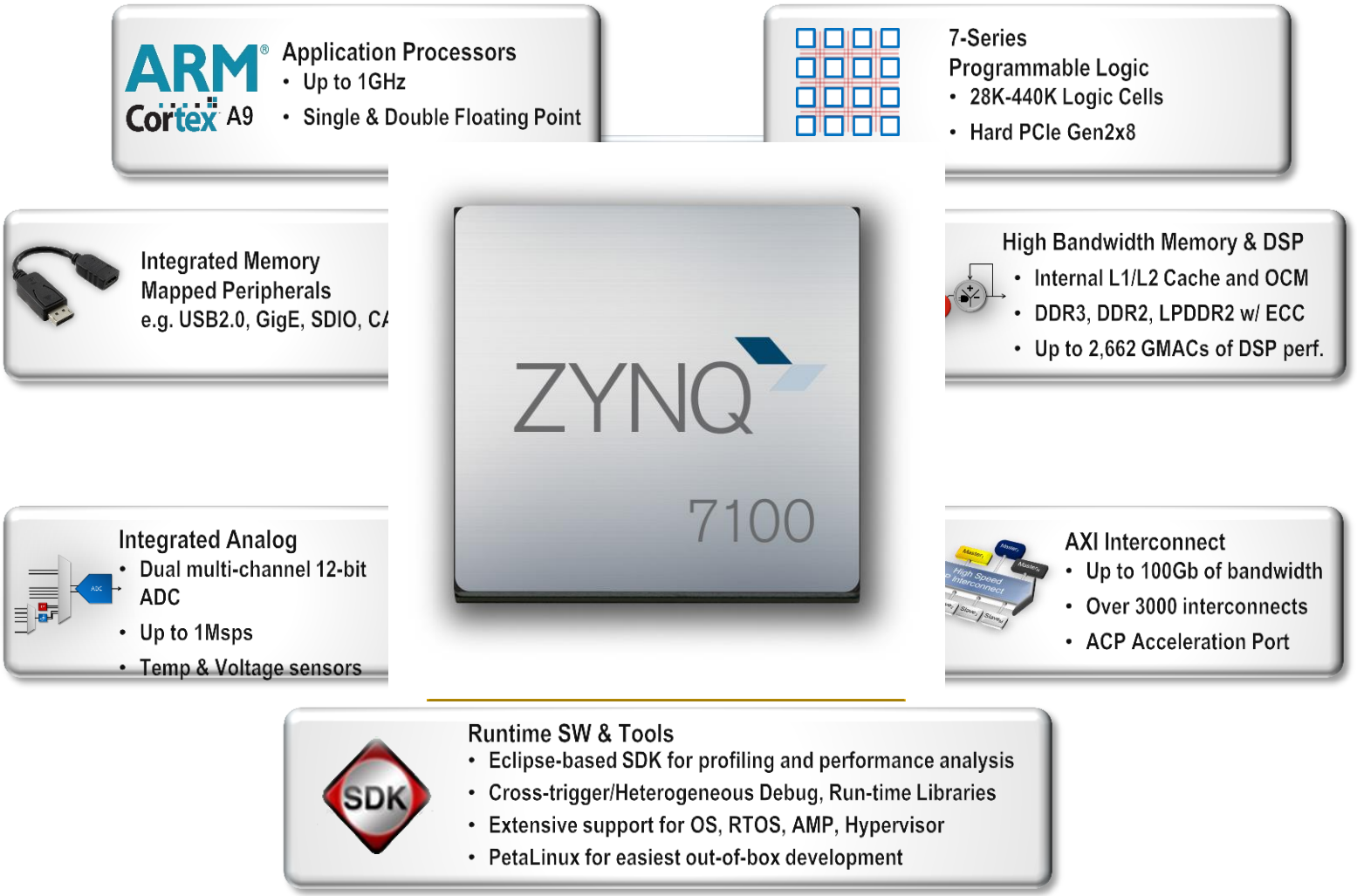
So, what's the problem?

- **What to do with all the functionality possible?**
- **Need programmability**
  - Not just hardware… but software programmability

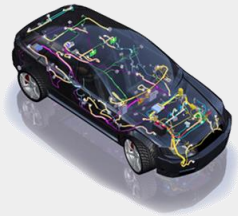


Electronics, Volume 38, Number 8, April 19, 1965

Enter Zynq 7000-AP SoC (2011)



Next generation challenges



- Power, Performance, Cost drivers
- Power management
- 64bit processing
- Real-time processing
- Video and graphics processing
- Pervasive safety and security
- Higher levels of processor-fabric integration

Introducing the Zynq UltraScale+ MPSoC

ARM Cortex A53 & R5

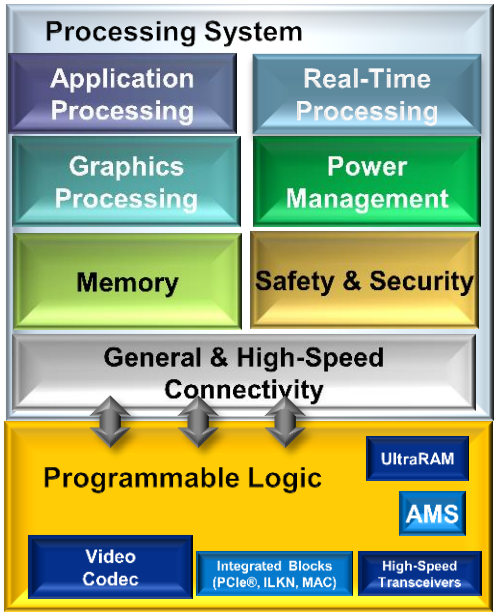
- Application processing subsystem
- Real Time processing subsystem
- Memory subsystem for max performance

IO, Video, Graphics

- Next-generation coherent interconnect
- High-speed I/Os (PCIE,USB3,SATA,GbE)
- Graphics and Video Processing Engines

Advanced Power Mgmt

- Fine-grained power reduction
- System-level software & run time opt



Security & Safety

- Enhanced Authentication, Encryption, Antitamper and trust
- Safety with industry standards support

Fabric Acceleration

- UltraScale+ fabric with time borrow
- FinFET performance and power
- HD UlltraRAM, and enhanced DSP

XCVRs & Protocols

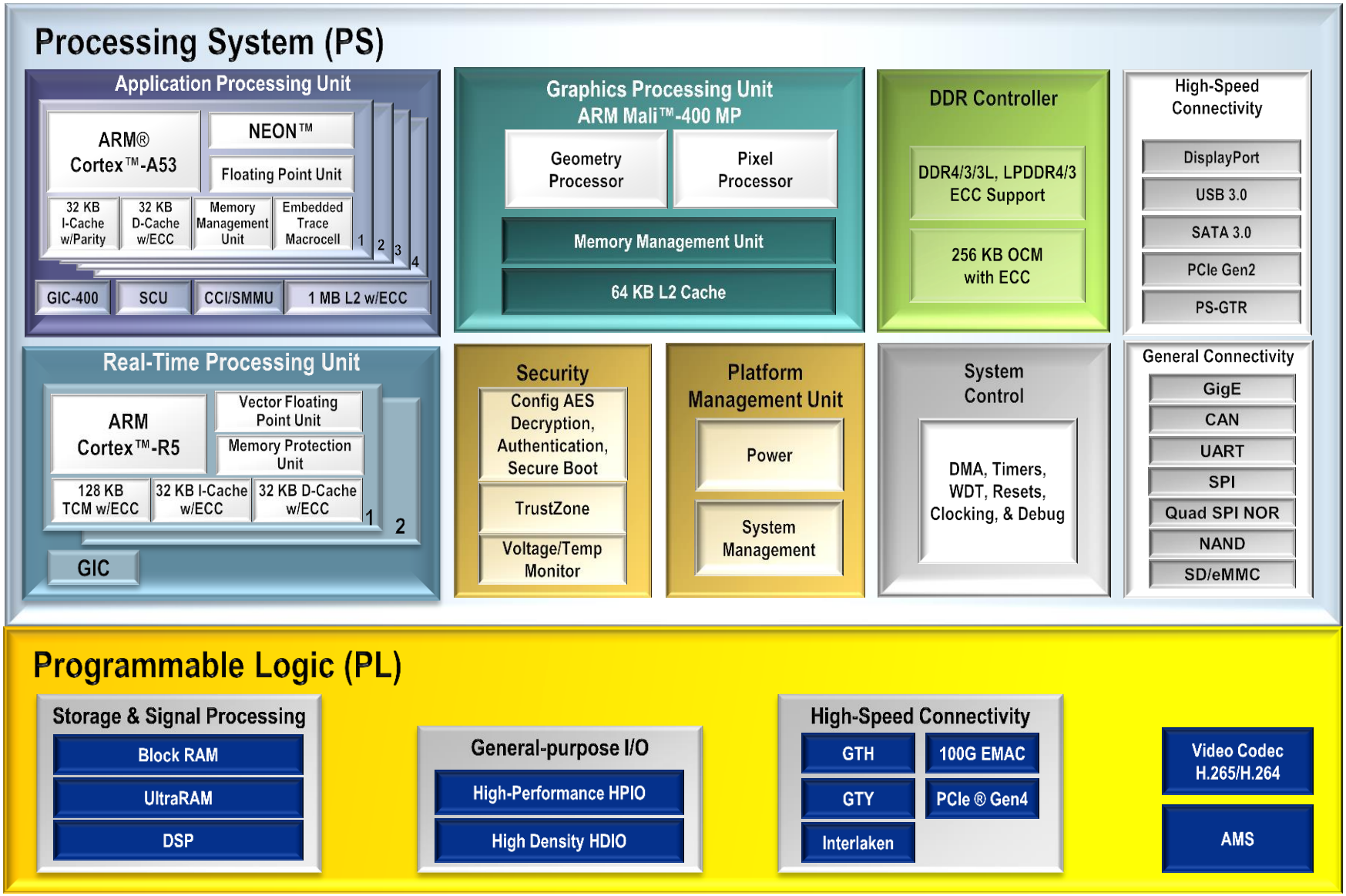
- Power efficient, 32Gbps
- 100G Ethernet and 150G Interlaken
- PCIe Gen3 & Gen4

Software & Tools

Run Time (Xilinx)	Run Time (Ecosystem)	Tools
- Linux (64b)	- FreeRTOS	- Xilinx SDK
- Hypervisor	- Micrium	- Vivado®
- OpenAMP	- WindRiver & More	- SDx environments

Zynq UltraScale+ MPSoC Overview

Heterogeneous Multi-Processing



Introducing the Zynq UltraScale+ MPSoC

ARM Cortex A53 & R5

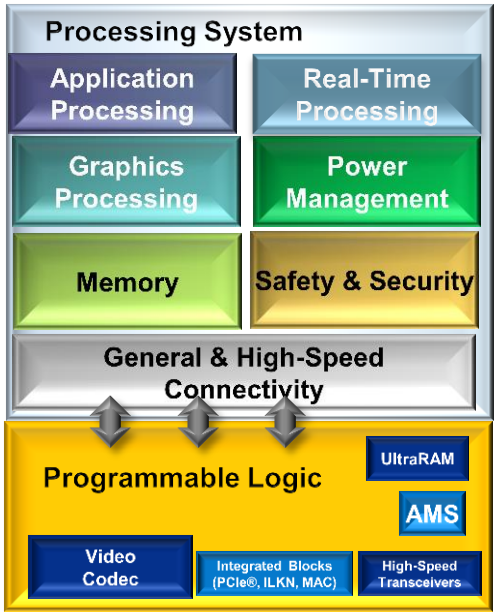
- Application processing subsystem
- Real Time processing subsystem
- Memory subsystem for max performance

IO, Video, Graphics

- Next-generation coherent interconnect
- High-speed I/Os (PCIE,USB3,SATA,GbE)
- Graphics and Video Processing Engines

Advanced Power Mgmt

- Fine-grained power reduction
- System-level software & run time opt

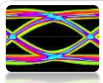


Security & Safety

- Enhanced Authentication, Encryption, Antitamper and trust
- Safety with industry standards support

Fabric Acceleration

- UltraScale+ fabric with time borrow
- FinFET performance and power
- HD UlltraRAM, and enhanced DSP



XCVRs & Protocols

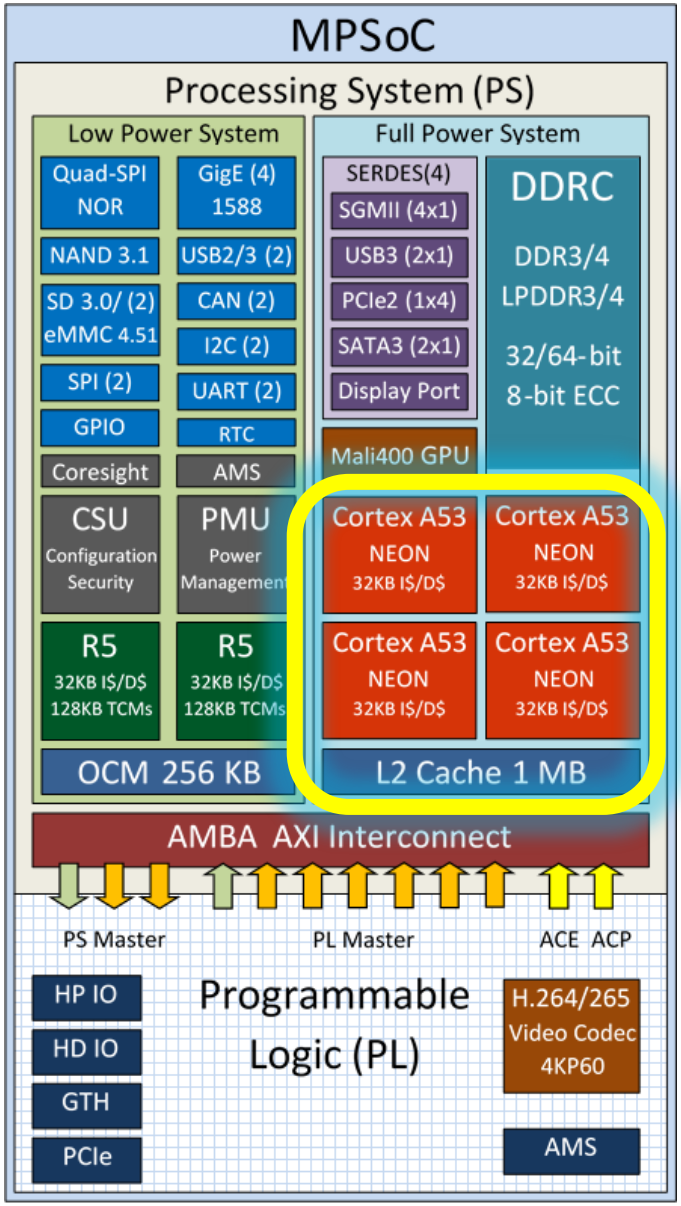
- Power efficient, 32Gbps
- 100G Ethernet and 150G Interlaken
- PCIe Gen3 & Gen4

Software & Tools

Run Time (Xilinx)	Run Time (Ecosystem)	Tools
- Linux (64b)	- FreeRTOS	- Xilinx SDK
- Hypervisor	- Micrium	- Vivado®
- OpenAMP	- WindRiver & More	- SDx environments

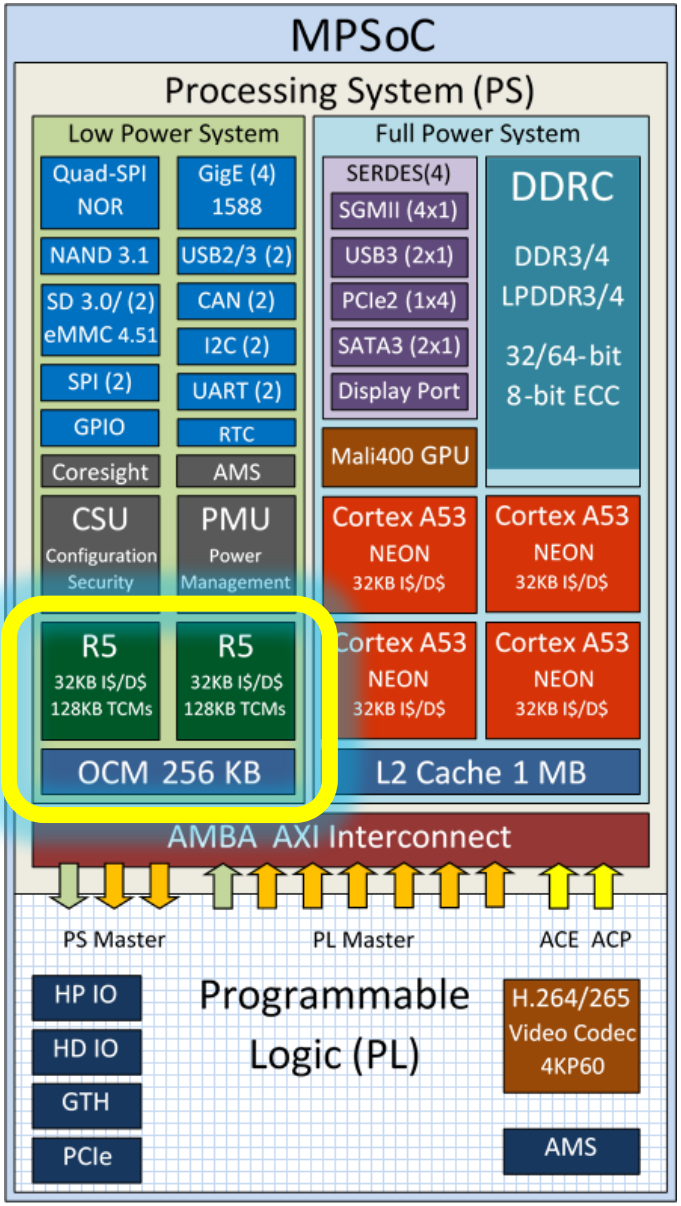
Application Processing Subsystem

- **Quad Cortex-A53 64-bit CPU**
  - 32KB each of L1 I & D$ with ECC/Parity
  - 1 MB L2 Cache with ECC
  - Virtualization Support
  - Crypto instructions support
- **IO Coherency with ACP & ACE-Lite**
- **Full Coherency between APU & PL**
- **Up to 1.5 GHz Frequency**
- **Power-gating**
  - Per core power-gating
  - L2 power-gating



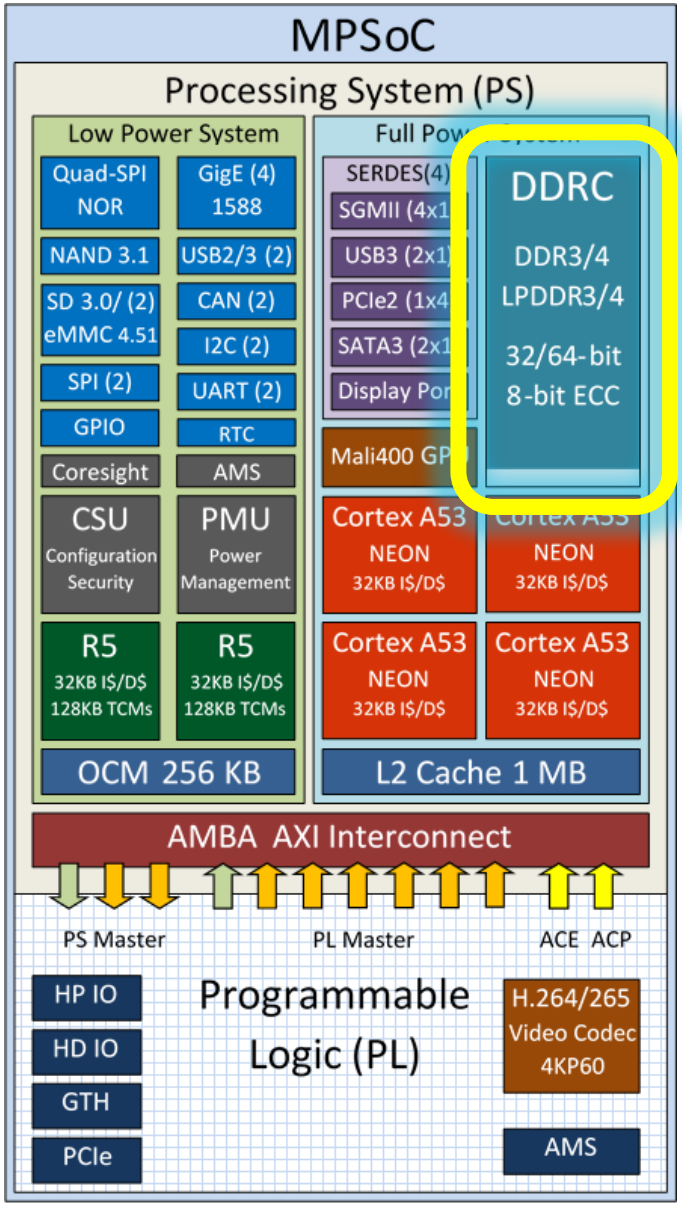
Real-time Processing Subsystem

- **Dual Core Cortex-R5 RPU**
  - Cortex-R5 Lockstep
  - Single & Double precision FPU
  - 32KB of L1 I & D caches w/ ECC
  - 256KB TCM with ECC
- **OCM (On-Chip-Memory)**
  - 256KB OCM with ECC
  - AXI Exclusive monitor support
  - Can be partitioned between different subsystems
- **Up to 600MHz Frequency**
- **Low power domain**
  - Full power domain completely powered off
  - R5s & USBs power-gate-able



Memory Subsystem

- **Six-port DDR Controller**
  - Supports exclusive monitors
- **32 or 64-bit DDR with ECC**
- **DDR3/4 and LPDDR3/4**
- **Up to 2400Mbps**
- **QoS support for 3 traffic classes**
  - Low latency, Real-time, Best-effort
  - Guaranteed latency for RT
- **Memory protection, partitioning, and TrustZone support**
  - Using XMPU



Introducing the Zynq UltraScale+ MPSoC

ARM Cortex A53 & R5

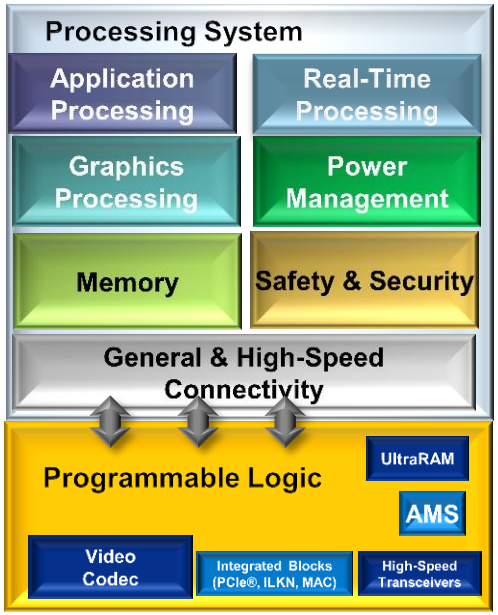
- Application processing subsystem
- Real Time processing subsystem
- Memory subsystem for max performance

IO, Video, Graphics

- Next-generation coherent interconnect
- High-speed I/Os (PCIE,USB3,SATA,GbE)
- Graphics and Video Processing Engines

Advanced Power Mgmt

- Fine-grained power reduction
- System-level software & run time opt



Security & Safety

- Enhanced Authentication, Encryption, Antitamper and trust
- Safety with industry standards support

Fabric Acceleration

- UltraScale+ fabric with time borrow
- FinFET performance and power
- HD UlltraRAM, and enhanced DSP

XCVRs & Protocols

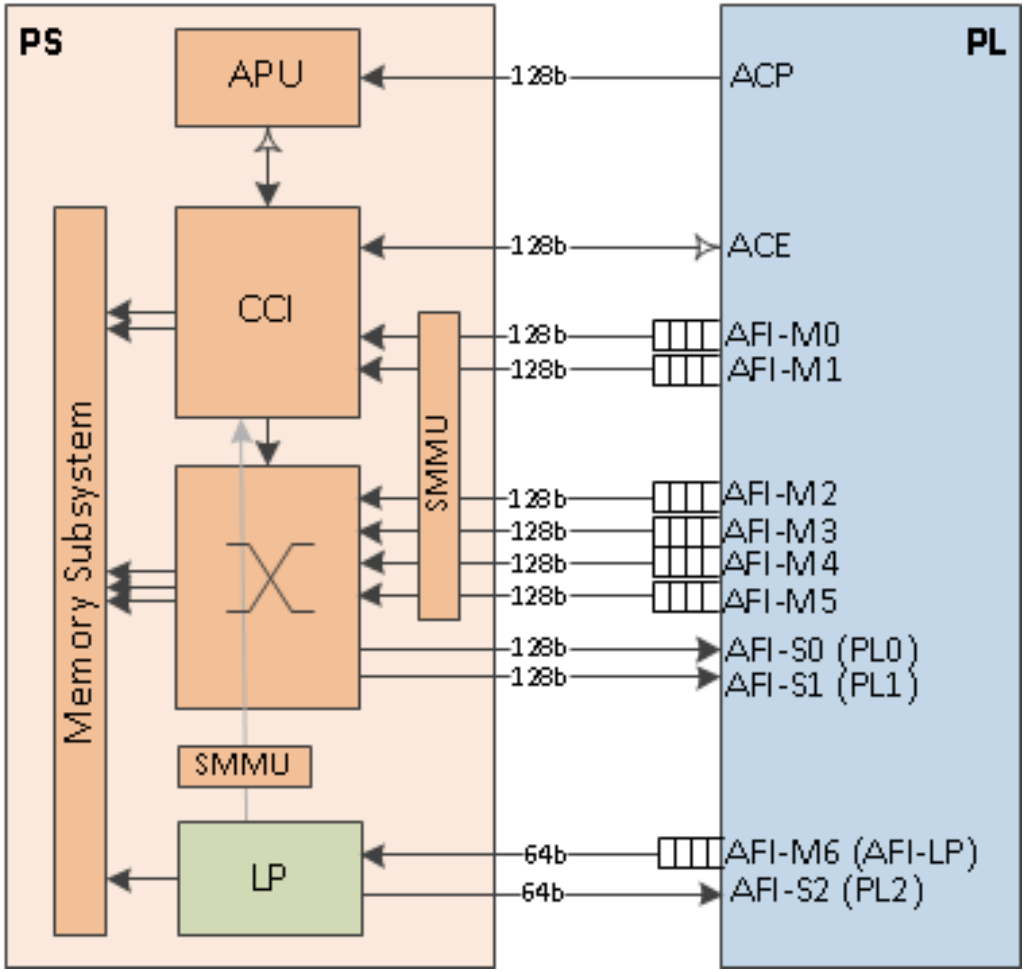
- Power efficient, 32Gbps
- 100G Ethernet and 150G Interlaken
- PCIe Gen3 & Gen4

Software & Tools

Run Time (Xilinx)	Run Time (Ecosystem)	Tools
- Linux (64b)	- FreeRTOS	- Xilinx SDK
- Hypervisor	- Micrium	- Vivado®
- OpenAMP	- WindRiver & More	- SDx environments

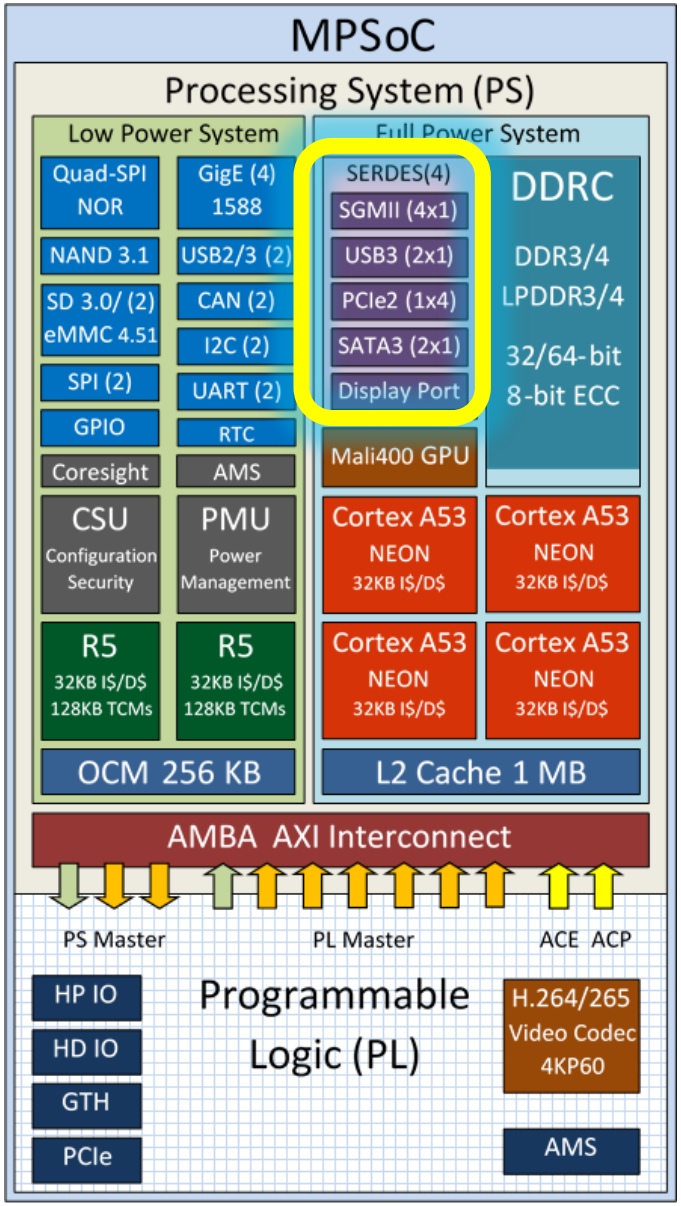
PS ⇔ PL “Data Mover” Interfaces

- **PL Master Ports**
  - AFI Master ports
  - PL IO-Coherency via CCI
  - PL virtualization via SMMU
- **PL Slave Ports**
  - PS to PL data movers
  - Memory mapped
- **ACP (Accelerated Coherency Port)**
  - ACP for IO (one-way) coherency
- **ACE (AXI Coherency Extn)**
  - Full coherency between PS & PL
- **Per-Port Bandwidth = 85Gbps**
  - Read+write bandwidth



High Speed I/Os

- **USB2/3**
  - 2 independent controllers
  - OTG, Host, Device
- **SATA3**
  - Up to 2 channels
- **Display Port**
  - 4KP30 support
  - 1-2 lanes
- **PCIe Gen2 Rootport or Endpoint**
  - PCIe Gen3/4 EP also hardened
- **SGMII for GbE**
  - 4 independent GbE controllers
- **Tightly integrated transceivers**



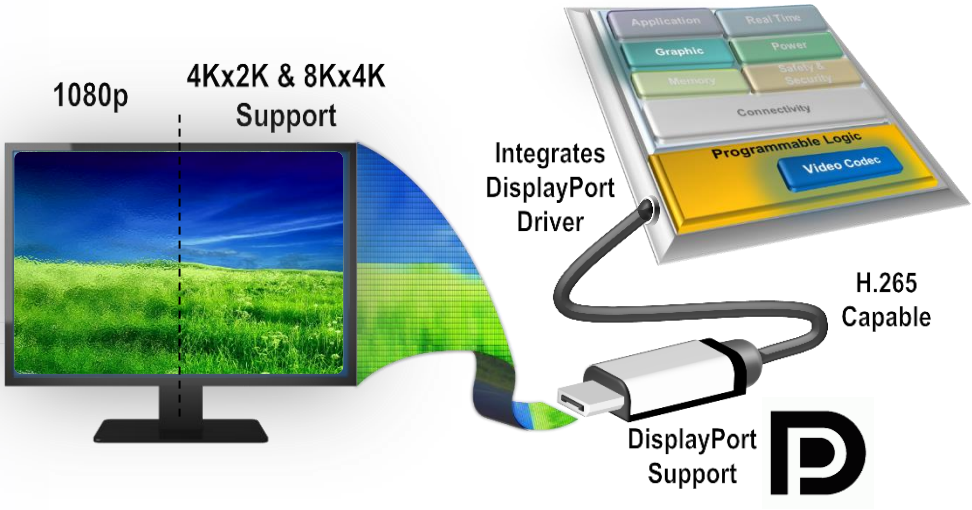
Dedicated Video Processing Engines

Graphics Processing Unit, Video Codec & DisplayPort

Video Codec Unit (VCU)

- More efficent vs. software implementation
  - Higher display density, faster encoding
  - Lower power consumption
- H.265 (HEVC) 8Kx4K (15 fps) 4Kx2K (60 fps)
- 8 and 10 bit per color component
- I, P, B frame support for highest compression

Higher Performance & Lower Power Video Processing

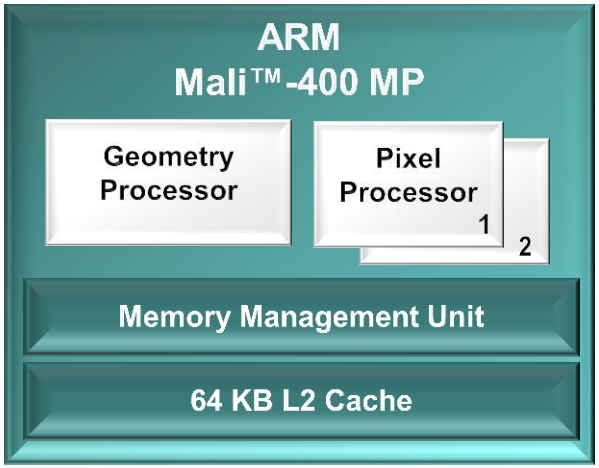


DisplayPort

- Video resolution up to 4Kx2K (30 fps)
- Audio up to 8 channels of 24-bit at up to 192 KHz
- Reducing BOM cost by eliminating display driver

Graphics Processing Unit (GPU)

- 3D visual, HMI, instrumentation, waveform display
- 1080p resolution graphics
- Mali-400 MP2 up to 667 MHz frequency



Introducing the Zynq UltraScale+ MPSoC

ARM Cortex A53 & R5

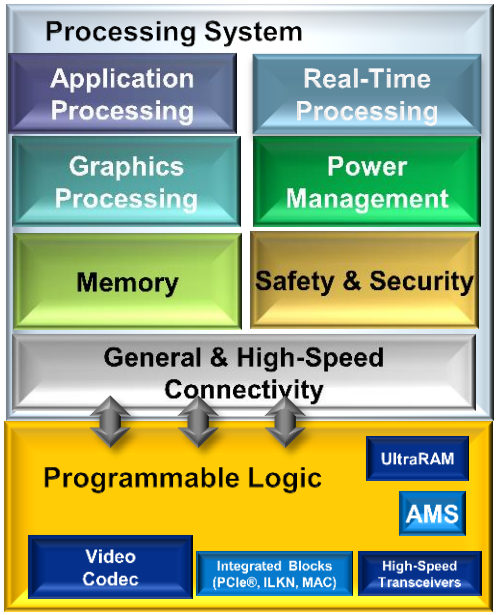
- Application processing subsystem
- Real Time processing subsystem
- Memory subsystem for max performance

IO, Video, Graphics

- Next-generation coherent interconnect
- High-speed I/Os (PCIE,USB3,SATA,GbE)
- Graphics and Video Processing Engines

Advanced Power Mgmt

- Fine-grained power reduction
- System-level software & run time opt



Security & Safety

- Enhanced Authentication, Encryption, Antitamper and trust
- Safety with industry standards support

Fabric Acceleration

- UltraScale+ fabric with time borrow
- FinFET performance and power
- HD UlltraRAM, and enhanced DSP

XCVRs & Protocols

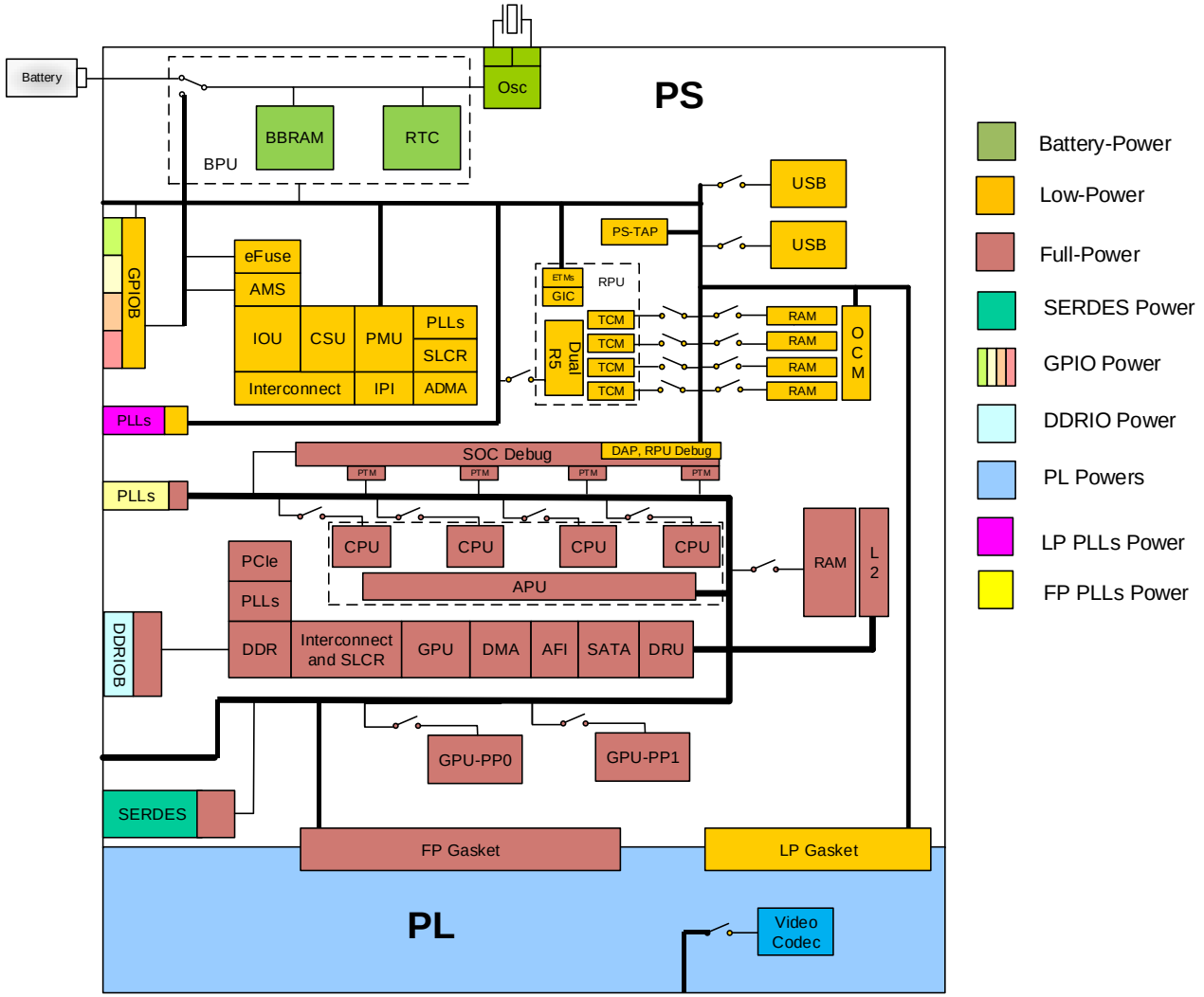
- Power efficient, 32Gbps
- 100G Ethernet and 150G Interlaken
- PCIe Gen3 & Gen4

Software & Tools

Run Time (Xilinx)	Run Time (Ecosystem)	Tools
- Linux (64b)	- FreeRTOS	- Xilinx SDK
- Hypervisor	- Micrium	- Vivado®
- OpenAMP	- WindRiver & More	- SDx environments

Power-Domains and Power-gating

- **Multiple power domains**
  - Low Power Domain
  - Full Power Domain
  - PL Power Domain
- **Power Gating**
  - A53 per core
  - L2 and OCM RAMs
  - GPU, USB
  - R5s & TCM
  - Video Codec
- **Sleep Mode**
  - 35mW sleep mode
  - Suspend to DDR with power off



Security, Safety & Reliability

Advanced Device-Level Secure Processing

- Information Assurance, Anti-Tamper, Trust
- Multi-layered Authentication for Secure System Boot
- Key Management & Revocation

Architected for Safe Systems

- IEC61508 & ISO26262 Functional Safety Standards
- Redundancy, Diversity and Lock-step
- Layered Partitioning: Core / Infrastructure / Peripherals

Delivering High Reliability

- High Availability Systems
- Error Detection & Handling
- Subsystem Isolation & Protection

Introducing the Zynq UltraScale+ MPSoC

ARM Cortex A53 & R5

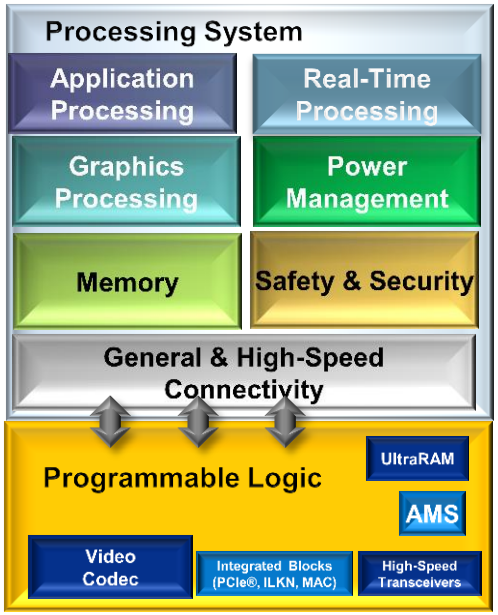
- Application processing subsystem
- Real Time processing subsystem
- Memory subsystem for max performance

IO, Video, Graphics

- Next-generation coherent interconnect
- High-speed I/Os (PCIE,USB3,SATA,GbE)
- Graphics and Video Processing Engines

Advanced Power Mgmt

- Fine-grained power reduction
- System-level software & run time opt



Security & Safety

- Enhanced Authentication, Encryption, Antitamper and trust
- Safety with industry standards support

Fabric Acceleration

- UltraScale+ fabric with time borrow
- FinFET performance and power
- HD UlltraRAM, and enhanced DSP

XCVRs & Protocols

- Power efficient, 32Gbps
- 100G Ethernet and 150G Interlaken
- PCIe Gen3 & Gen4

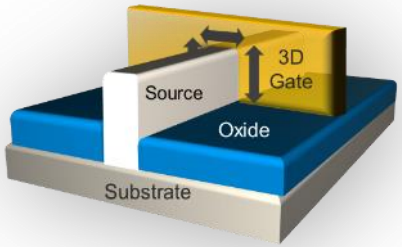
Software & Tools

Run Time (Xilinx)	Run Time (Ecosystem)	Tools
- Linux (64b)	- FreeRTOS	- Xilinx SDK
- Hypervisor	- Micrium	- Vivado®
- OpenAMP	- WindRiver & More	- SDx environments

Tuned Process for Optimal Performance/Watt

Optimal Operating Voltage Selection

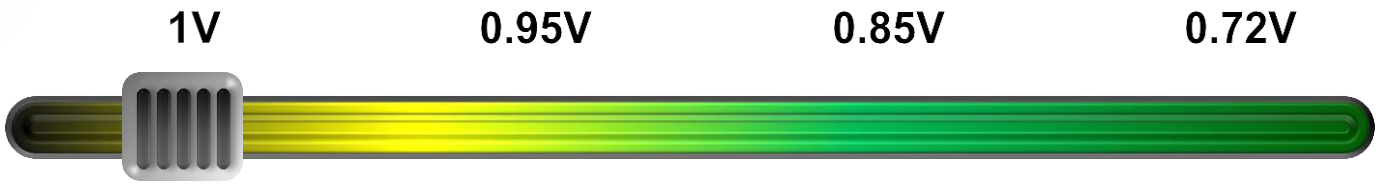
3D FinFET



3D Gate "wraps" around channel for more surface area, achieving

- ✓ Faster transistor on/off switching speeds for greater performance
- ✓ Lower leakage and operating voltage for lower power

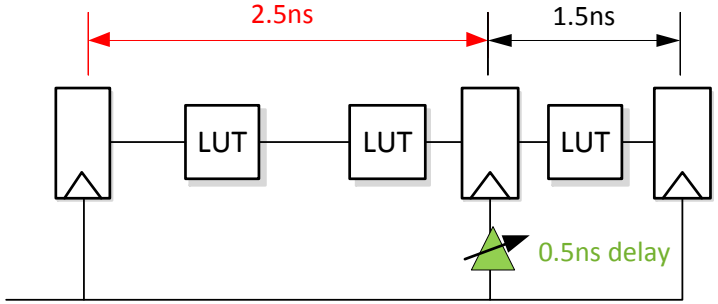
Tuned Operating Voltage (V_{CCINT})



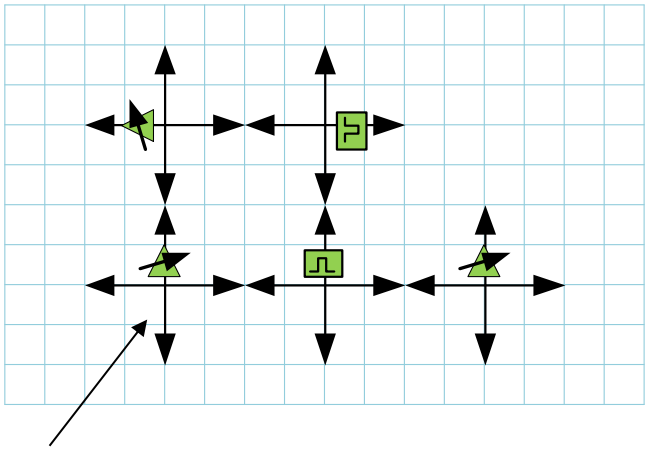
	7 Series (28nm) V_{NOM}	20% Greater Performance	UltraScale (20nm) V_{NOM}	33% Greater Performance	UltraScale+ (16nm) V_{NOM}	UltraScale+ (16nm) V_{LOW}
Normalized Fabric Performance	1.0x		1.2x		1.6x	1.2x
Normalized Total Power	1.0x		.7x		.8x	.5x
Performance/Watt	1.0x		1.7x		2x	2.4x

Time Borrow in the Fabric

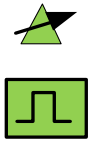
- **Time-borrowing concept**
  - Shift available slack from fast stages to performance-critical paths



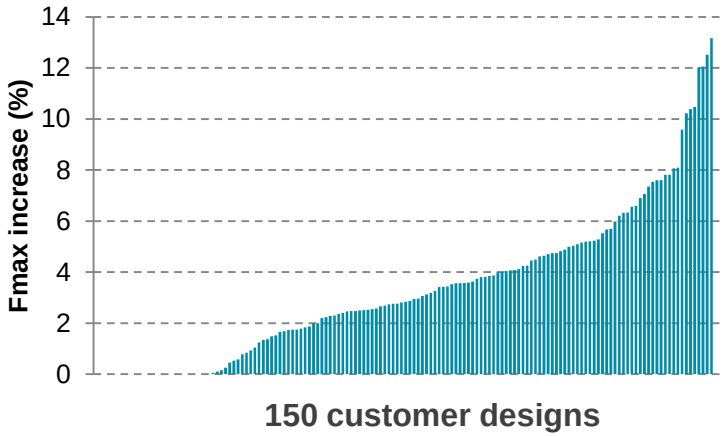
Example	Tmin	Fmax
Baseline	2.5 ns	400 MHz
Time Borrow	2 ns	500 MHz



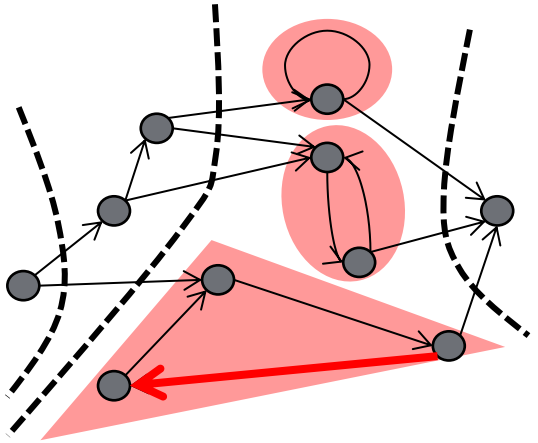
- **UltraScale+ time-borrowing platform**
  - Fine-grain delays to adjust clock skew
  - Programmable pulse generators for latch-based time borrow



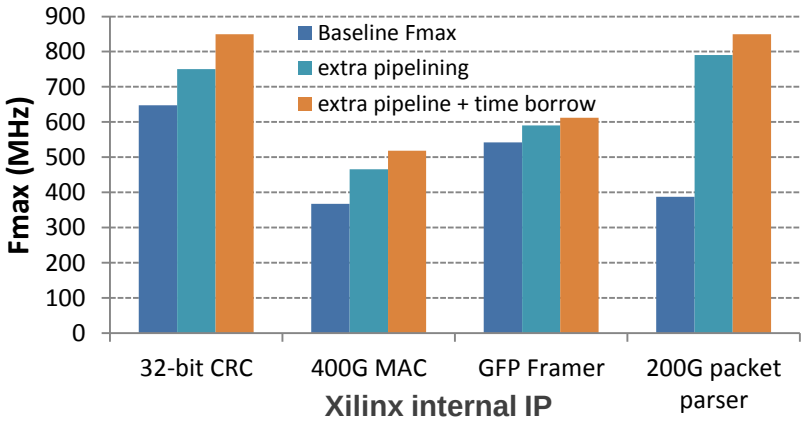
- **High Performance without design changes**
  - Very effective on high-performance designs
  - Transparent to customers, part of default flow



Extra Pipeline Analysis in Vivado



- **New automatic pipeline analysis**
  - Automatically analyses design at any SPR stage
  - Finds sequential loops, analyzes timing
  - Detailed latency vs Fmax analysis
- **Enables rapid design exploration**
  - Shows path to 800MHz+ Fmax
  - Suggests most efficient places to insert regs in RTL
- **Backward-compatible**
  - 7-series, UltraScale, UltraScale+
  - Compatible with HLS and SDAccel/SDSoC



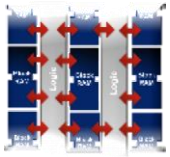
Clock	Added Latency	Ideal Fmax (MHz)	Ideal Delay (ns)	Requirement (ns)	WNS (ns)*	Added Pipe Reg	Total Pipe Reg
SYS_CLK	0	349.00	2.87	2.31	-0.56	n/a	0
SYS_CLK	1	354.69	2.82	2.31	-0.51	7693	7693
SYS_CLK	2	356.97	2.80	2.31	-0.49	5213	12906
SYS_CLK	3	364.78	2.74	2.31	-0.43	3613	16519
SYS_CLK	4	366.12	2.73	2.31	-0.42	7348	23867
SYS_CLK	5	537.24	1.86	2.31	0.45	20348	44215

UltraRAM: New Memory Technology

Distributed RAM
(bits to kilobits)

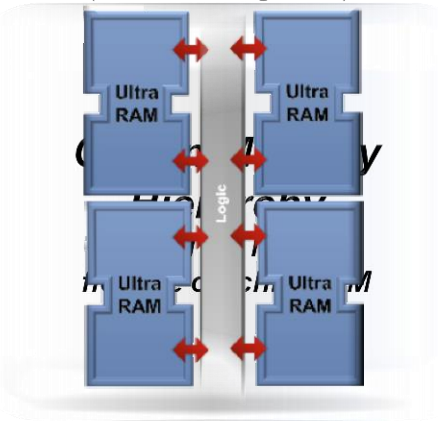
- Wide, shallow FIFOs
- Shift registers
- State machines

Block RAM
(10s of megabits)



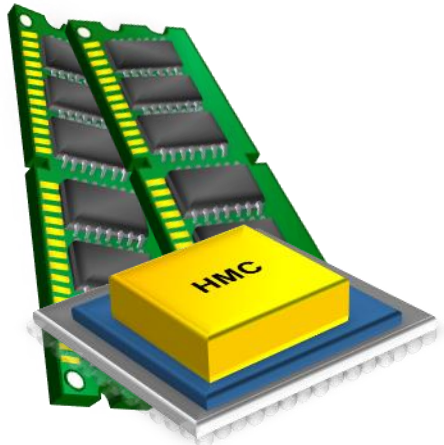
- Data/coefficient storage
- Deep FIFOs
- Shallow buffering

UltraRAM
(100s of megabits)



- Deep packet buffering
- Video buffering
- State, statistics, counters

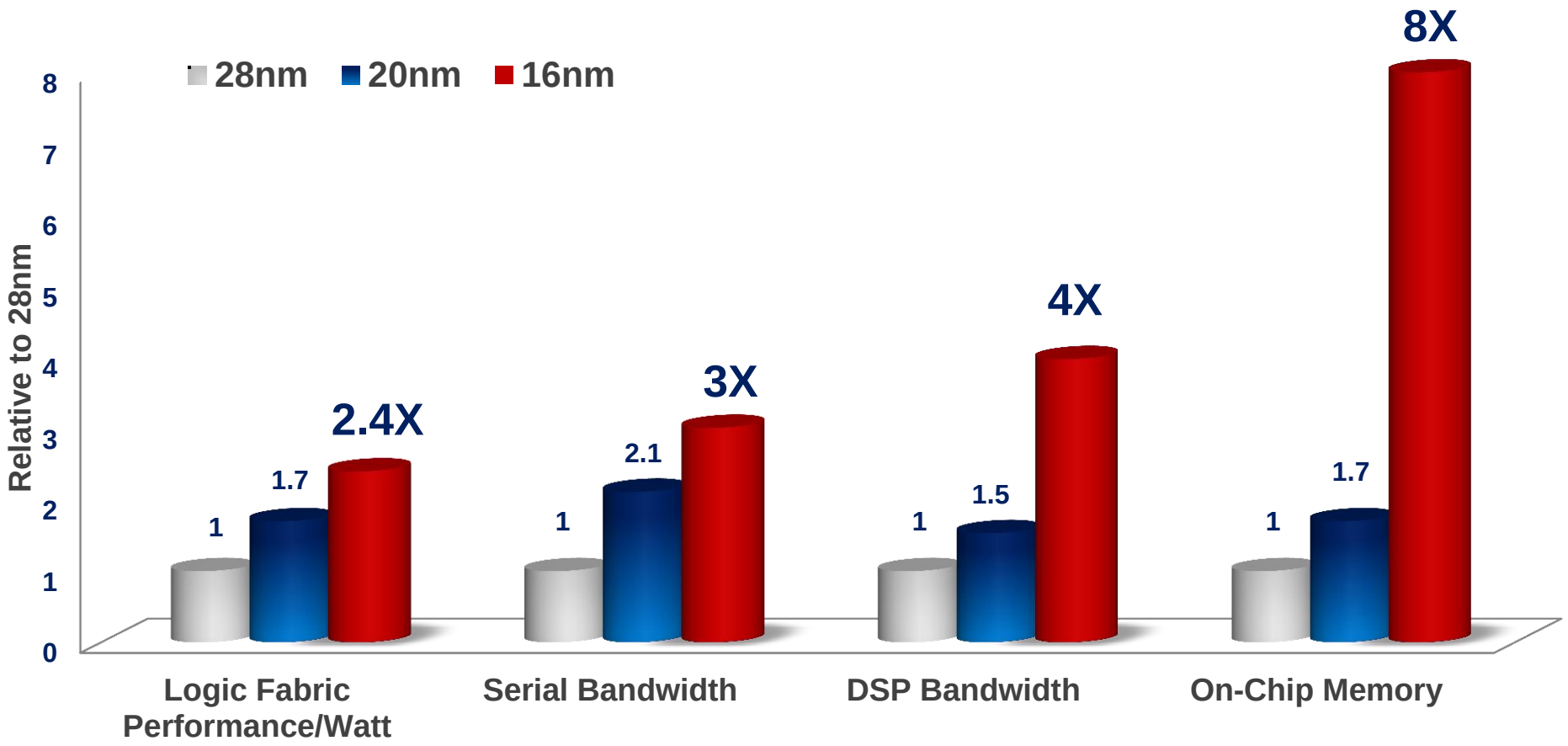
External Memory
(100s of megabits to gigabits)



Larger data storage

Up to 432 Mb to replace external memory for cost, power, performance

Unlocking Performance, Bandwidth, & Integration



Enhanced Fabric with FinFET performance	Up to 128 transceivers at up to 32.75 Gb/s	~12,000 DSP slices running at ~900 MHz	UltraRAM for SRAM device replacement

Introducing the Zynq UltraScale+ MPSoC

ARM Cortex A53 & R5

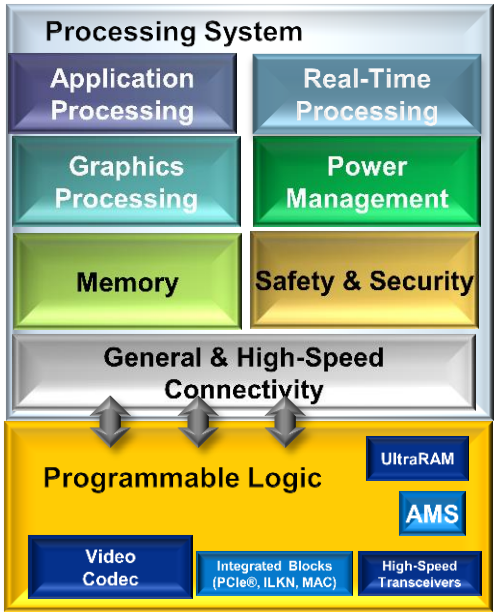
- Application processing subsystem
- Real Time processing subsystem
- Memory subsystem for max performance

IO, Video, Graphics

- Next-generation coherent interconnect
- High-speed I/Os (PCIE,USB3,SATA,GbE)
- Graphics and Video Processing Engines

Advanced Power Mgmt

- Fine-grained power reduction
- System-level software & run time opt



Security & Safety

- Enhanced Authentication, Encryption, Antitamper and trust
- Safety with industry standards support

Fabric Acceleration

- UltraScale+ fabric with time borrow
- FinFET performance and power
- HD UlltraRAM, and enhanced DSP

XCVRs & Protocols

- Power efficient, 32Gbps
- 100G Ethernet and 150G Interlaken
- PCIe Gen3 & Gen4

Software & Tools

Run Time (Xilinx)	Run Time (Ecosystem)	Tools
- Linux (64b)	- FreeRTOS	- Xilinx SDK
- Hypervisor	- Micrium	- Vivado®
- OpenAMP	- WindRiver & More	- SDx environments

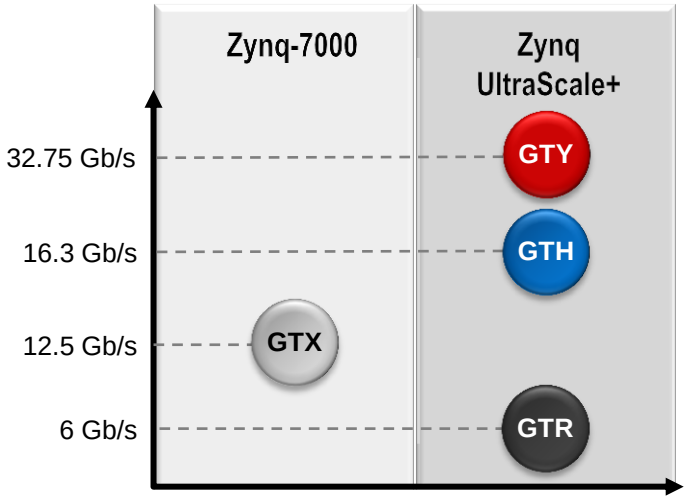
Enhanced transceivers

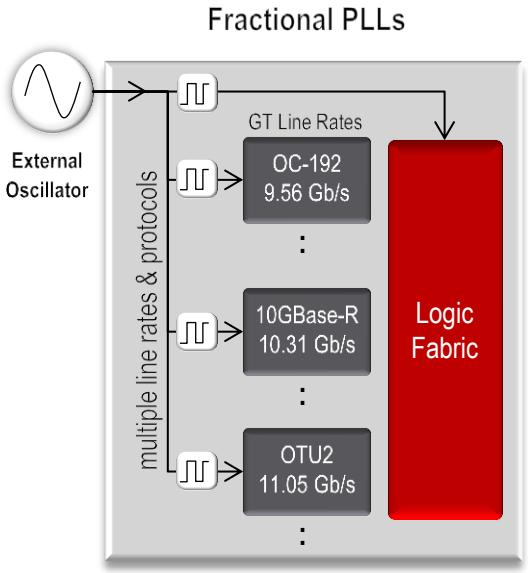
Diverse, Power Efficient SerDes for Bandwidth

- 16G (GTH) & 32G (GTY) transceivers in PL,
- 6G (GTR) in PS for direct access to key processing elements, with full PHY/IP compliance for key protocols:
  - USB, SATA, DisplayPort, PCIe, Ethernet

Fractional PLLs to Reduce BOM Cost

- Single external oscillator generates GT & logic fabric clocks for multiple non-integer line rates
- Available in GTH, and GTY transceivers





Introducing the Zynq UltraScale+ MPSoC

ARM Cortex A53 & R5

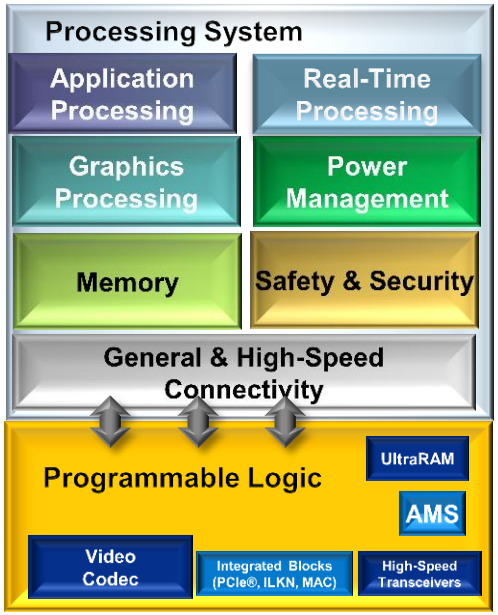
- Application processing subsystem
- Real Time processing subsystem
- Memory subsystem for max performance

IO, Video, Graphics

- Next-generation coherent interconnect
- High-speed I/Os (PCIE,USB3,SATA,GbE)
- Graphics and Video Processing Engines

Advanced Power Mgmt

- Fine-grained power reduction
- System-level software & run time opt



Security & Safety

- Enhanced Authentication, Encryption, Antitamper and trust
- Safety with industry standards support

Fabric Acceleration

- UltraScale+ fabric with time borrow
- FinFET performance and power
- HD UlltraRAM, and enhanced DSP

XCVRs & Protocols

- Power efficient, 32Gbps
- 100G Ethernet and 150G Interlaken
- PCIe Gen3 & Gen4

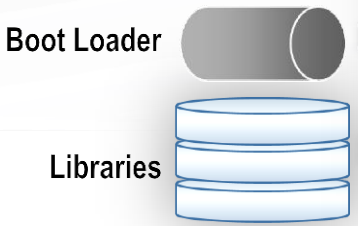
Software & Tools

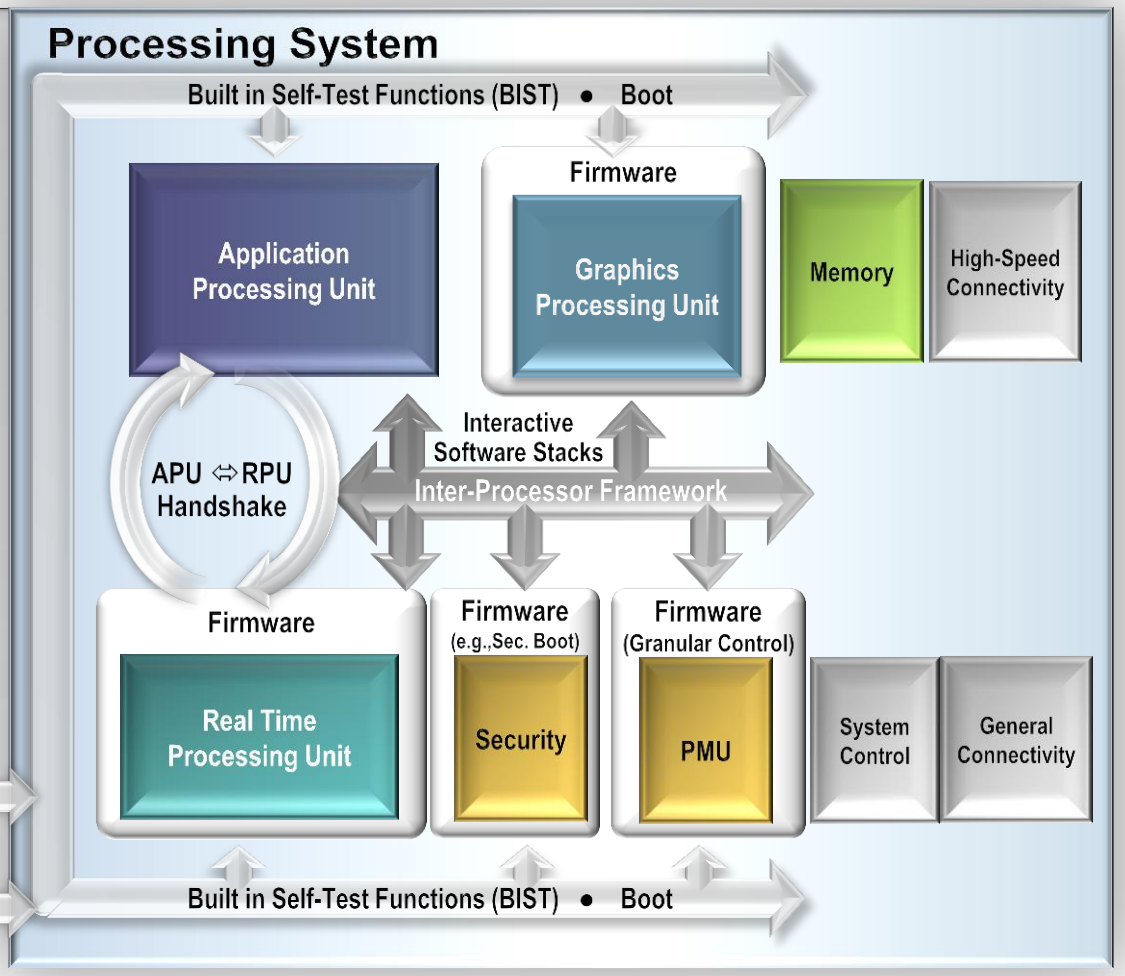
Run Time (Xilinx)	Run Time (Ecosystem)	Tools
- Linux (64b)	- FreeRTOS	- Xilinx SDK
- Hypervisor	- Micrium	- Vivado®
- OpenAMP	- WindRiver & More	- SDx environments

System Software

System Software for MPSoC

- Security Firmware
  - Decryption, authentication, sec. boot
- FSBL, uBoot
- ARM® Trusted Firmware
- Software Test Libraries (BIST)
- Power Management Firmware
  - Granular control of resources
- Inter-Processor Framework - OpenAMP



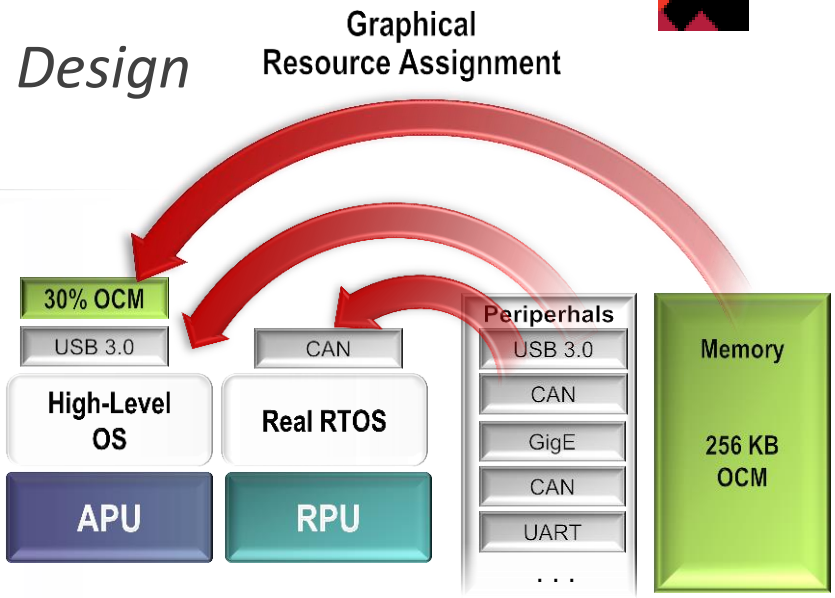


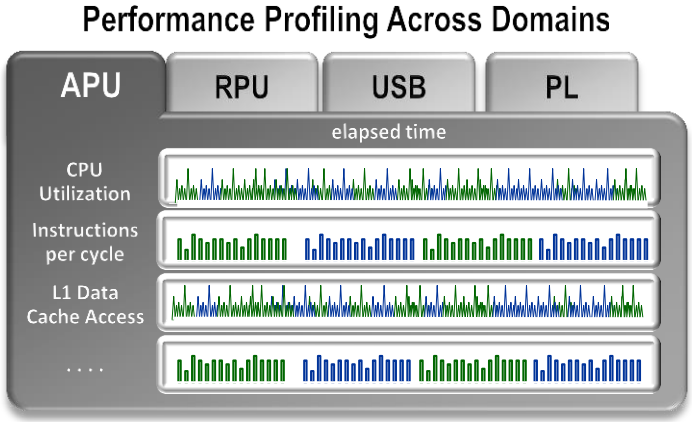
Xilinx SDK

Integrated Environment for *Software Design*

New & Enhanced Software Tools

- Heterogeneous Multicore Debug
  - Debug & cross triggering for APU/RPU/MicroBlaze™ Processor
- System-level Profiling and Performance Analysis Tools
  - Analysis for interfaces
  - Across processing & Programmable Logic (PL) domains
- Multi-OS Boot Image Tool
  - Creates boot image(s)
  - Supports output from Xilinx & 3rd party IDEs
- System Resource Partitioning Tool
  - Graphic assignment of resources





Industry-standard Tools Support

- Enabling developer-preferred dev /debug environments

YOKOGAWA

ARM DS Development Tools

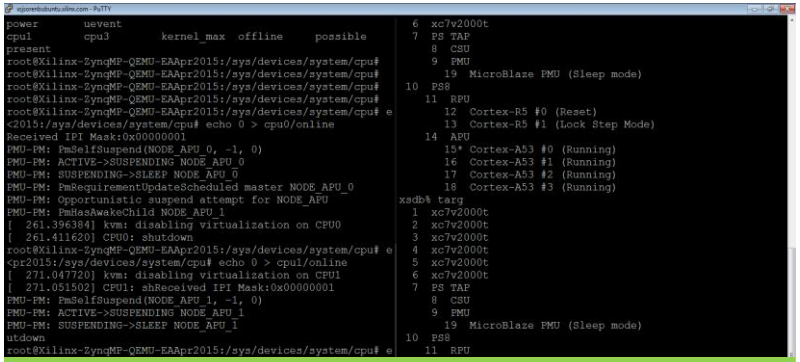
LAUTERBACH DEVELOPMENT TOOLS

KMC

XILINX ALL PROGRAMMABLE

Software and Systems running on platforms

Linux CPU Hotplug on REmuS



- REmuS (below) runs Linux, ARM Trusted Firmware, and PMU Firmware
- User uses hotplug interface to power down A53 CPUs

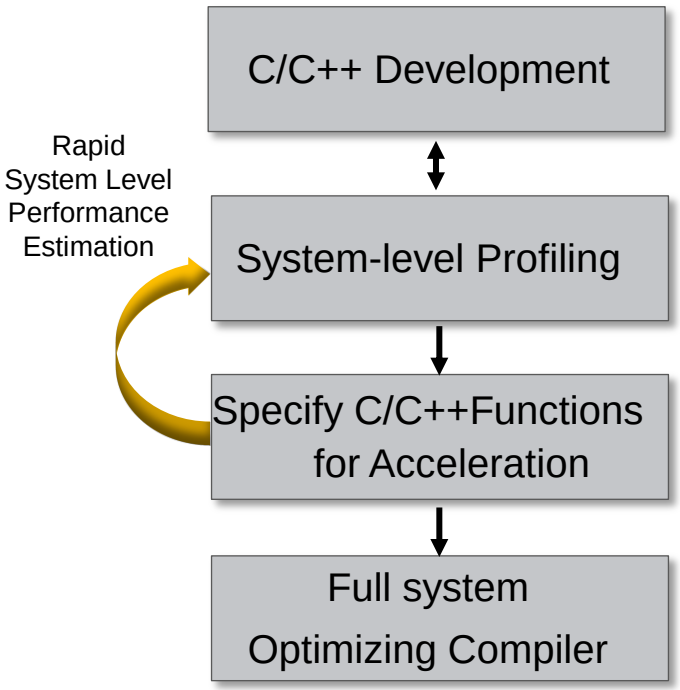
Doom on QEMU and REmuS

- **Virtualization Passthrough**
  - Linux guest using a GEM
  - SMMU used in QEMU
- **XEN tools**
  - Built packages with Yocto
  - Integrated into PetaLinux
- **ParaVirtual FrameBuffer**
  - VNC over PV FrameBuffer
  - DomU runs doom (playable)

SDSoC Development Environment

C/C++ Programming for Zynq SoC and MPSoC

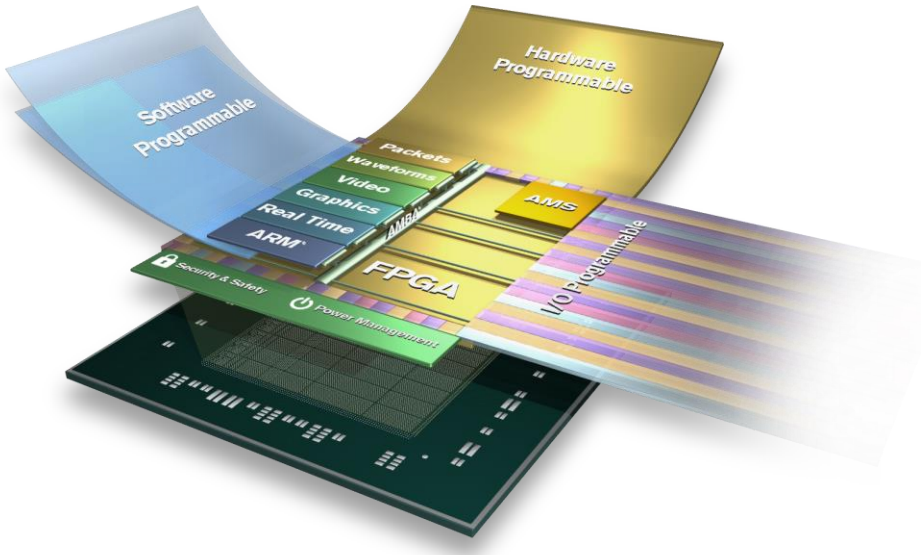
SDSoC™ Environment



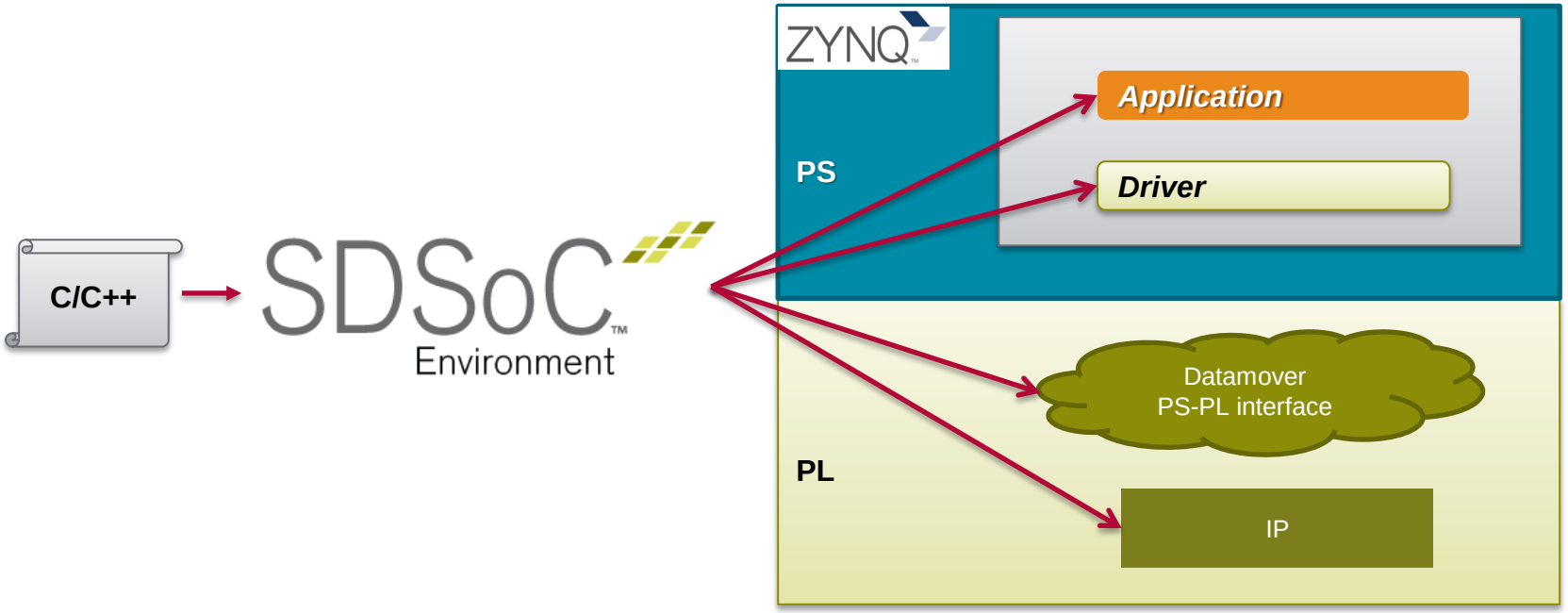
ZYNQ® SoC

ZYNQ® MPSoC

- ASSP-like programming experience
- System-level profiling
- Full system optimizing compiler
- Expert use model for platform developers and system architects

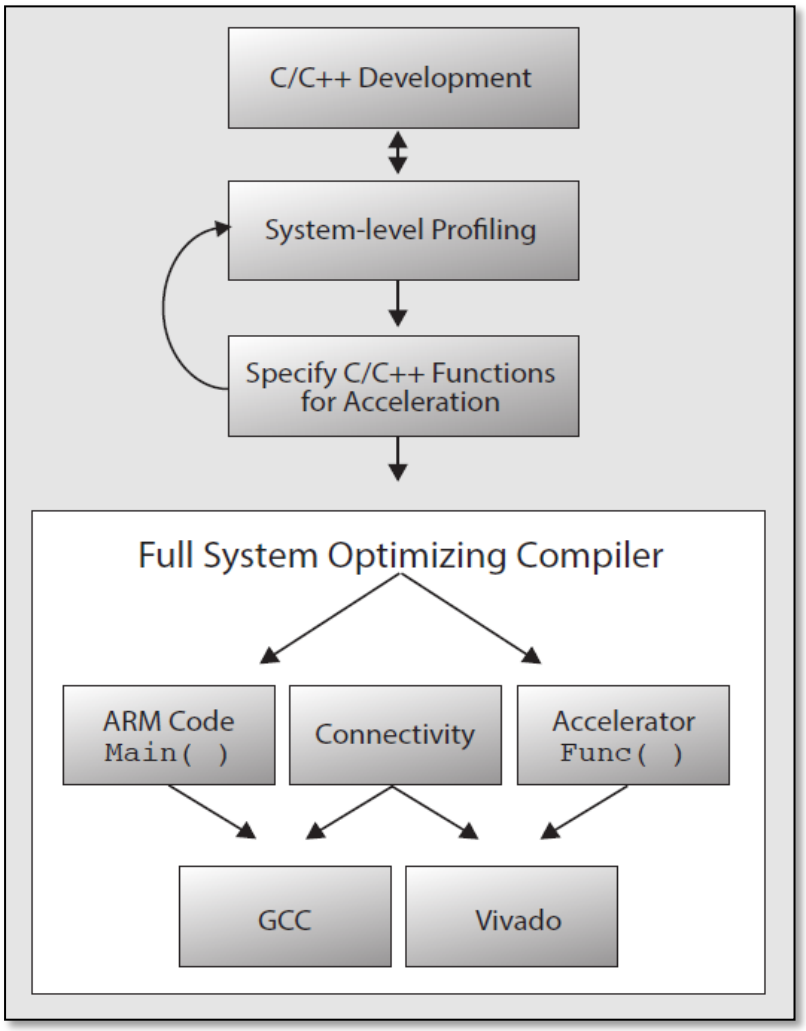


SDSoC: ASSP Like Programming Experience



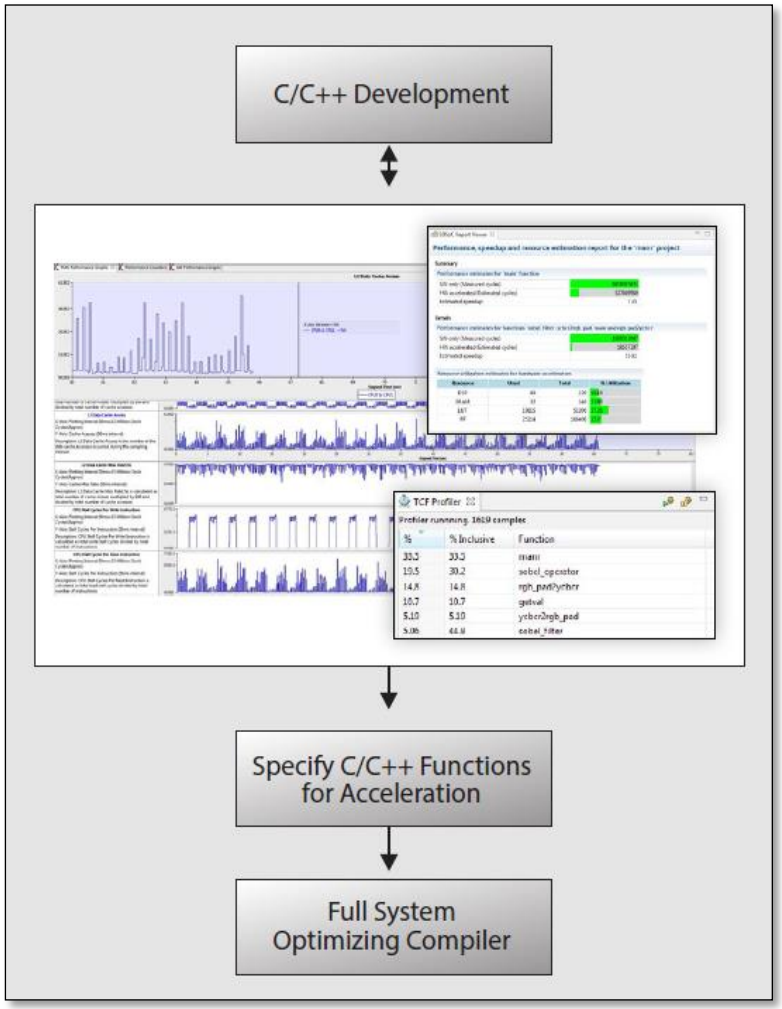
- C/C++ application to a complete system using easy-to-use Eclipse IDE
- 'One-click' function acceleration in Programmable Logic
- Eliminates manual configuration of SW drivers and HW connectivity
- Support for multiple OSes and optimized libraries
- Enables reuse of optimized HDL IP Blocks as C-callable libraries

SDSoC: Full System Optimizing Compiler



- **Automated Connectivity Optimization**
  - Finds the Data Mover and PS-PL interface for optimal dataflow
  - Rapid exploration of different system connectivity topologies
- **Rapid Software Configurable Application Acceleration using C/C++**
  - Automated function acceleration in programmable logic
  - Up to 100X increase in performance vs. software
  - System optimized for latency, bandwidth, and hardware utilization

SDSoC: System Level Profiling



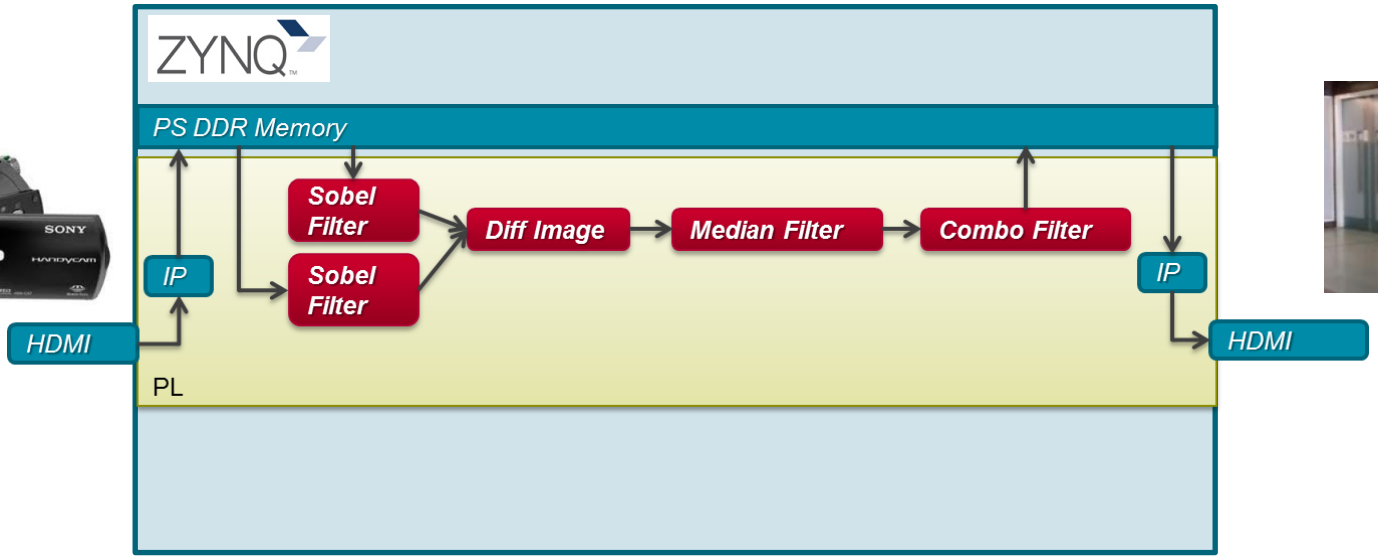
- **Rapid system performance estimation**
  - Full system estimation (programmable logic, data communication, processing system)
  - Reports SW/HW cycle level performance and hardware utilization
- **Automated performance measurement**
  - Runtime measurement by instrumentation of cache, memory, and bus utilization

1080p60 Motion Detection using SDSoC

```
main(){
  get_camera(A);
  sobel(A,B);
  diff(B,C);
  …
  display(out);
}
```

Video Platform

SDSoC™ Environment



C/C++ programming for Zynq – 2 weeks with multiple iterations

Summary

- **Zynq UltraScale+ MPSoC: 2nd Generation SoC from Xilinx**
  - Applications processing, Real-time, Graphics, Video, Serial connectivity
  - Power management, Safety, Security
  - SDSoC: Full system optimizing compiler
- **More than Moore: Architectural innovation**
  - 3x CPU performance and 4.5x memory bandwidth (SoC)
  - UltraScale+ fabric: 60% higher performance, 2.5x performance/watt (FPGA)
  - 3rd generation of silicon interposer technology (3D IC)
- **Taped out in Jun 2015 on TSMC 16FF+**
  - Significant power and performance benefits with 3D FinFet transistors
  - Diverse SW and systems running on multiple platforms today